




Multivariate analysis of morphological diversity among closely related *Daucus* species and subspecies in Tunisia

Najla Mezghani  · Jihene Ben Amor · David M. Spooner · Phillip W. Simon · Neila Mezghani · Hiba Boubaker · Ahmed M'rad Namji · Slim Rouz · Cherif Hannachi · Mohamed Neffati · Neji Tarchoun

Received: 19 October 2016 / Accepted: 27 February 2017
© Springer Science+Business Media Dordrecht 2017

Abstract The genus *Daucus* includes about 20–25 species worldwide. Northern Africa represents a major center of diversity of *Daucus*, with Tunisia thought to contain 11 species and seven subspecies. The greatest taxonomic problems, however, and the greatest economic importance relative to immediate use in breeding, is in a group of species and subspecies in the *Daucus carota* L. clade, all containing $2n = 18$ chromosomes. We assessed morphological diversity from a *Daucus* L. germplasm collection of nine individuals each from 45 accessions (405 individuals in total), at the National Gene Bank of Tunisia, of the following wild and one cultivated members of this clade: *D. carota* subsp. *capillifolius* (Gilli) Arbizu, *D. carota* subsp. *carota*, *D. carota* subsp. *gummifer* (Syme) Hook. f., *D. carota* subsp. *sativus* (Hoffm.) Arcangeli (cultivated), *D. sahariensis* Murb., and putative hybrids of *D. carota* subsp. *carota* and subsp.

capillifolius. A prior study showed the effectiveness of fruit characters to identify several species and subspecies in the collection, but distinction between some closely related *D. carota* subspecies was difficult. In order to resolve the taxonomic classification, we tested 32 quantitative and qualitative morphological characters from leaves, stems and flowers on a field collection of 45 accessions corresponding to the different species/subspecies. The Shannon–Weaver diversity (H') index was used to study the phenotypic diversity. The estimated H' ranged from monomorphic for umbel type, position of involucre bracts on primary umbel, anther color, and symmetry of peripheral flowers to highly polymorphic for other traits. The highest (0.98) and the lowest (0.26) H' values were recorded for flowering pattern within plants and foliage coverage. Multivariate analyses of principal components and dendrograms of all data and

N. Mezghani (✉)
Banque Nationale de Gènes, Boulevard du Leader Yasser Arafat, Z. I Charguia 1, 1080 Tunis, Tunisia
e-mail: najla_mezghani@yahoo.fr

J. Ben Amor · H. Boubaker · A. M'rad Namji · C. Hannachi · N. Tarchoun
Institut Supérieur Agronomique de Chott Mariem, 4042 Sousse, Tunisia

D. M. Spooner · P. W. Simon
USDA-ARS, Vegetable Crops Research Unit, Department of Horticulture, University of Wisconsin, 1575 Linden Drive, Madison, WI 53706-1590, USA

N. Mezghani
Centre de recherche LICEF, TELUQ university, 5800, rue Saint-Denis, bureau 1105, Montréal (Québec) H2S 3L5, Canada

S. Rouz
Institut Supérieur Agronomique de Mograne, 1121 Zaghouan, Tunisia

M. Neffati
Institut des Régions Arides, 4119 Medenine, Tunisia

canonical discriminate analysis of the quantitative data supported the subdivision of the *Daucus* collection into five groups with various degrees of distinctness: (1) *D. sahariensis*, (2) *D. carota* subsp. *capillifolius*, (3) *D. carota* subsp. *carota*, very similar to (4) *D. carota* subsp. *gummifer*, and (5) *D. carota* subsp. *sativus* intergrading with putative hybrids between *D. carota* subsp. *capillifolius* and *D. carota* subsp. *carota*. Individual character state distribution plots provide useful characters and insights into taxonomic problems in the *D. carota* clade that we here discuss in reference to ongoing molecular studies in *Daucus*.

Keywords *Daucus* · Multivariate analysis · Subspecies · Taxonomy

Abbreviations

NCRPIS North Central Regional Plant Introduction Station
 NGBT National Gene Bank of Tunisia

Introduction

Gene banks play a critical role in the conservation of agro-biodiversity at national and international levels (Rao et al. 2006; Agrawal et al. 2007). Germplasm of diverse plant species is maintained in gene banks worldwide, with individual collections holding anywhere from hundreds to tens of thousands of accessions (FAO 1997). The conservation, management, and use of germplasm maintained in gene banks pose many challenges genetic resources conservators and researchers using the collections. Central to sustainable conservation and effective use of the collections for breeding and other research is the knowledge of the genetic diversity and taxonomy present in the gene bank. Hence, the characterization of the accessions maintained in the collection and the examination of the genetic relationship between them is important for the efficient conservation and use of the collections (Teklu et al. 2006).

Daucus carota L. is a morphologically diverse species found in wild or feral form throughout the Mediterranean, southwest Asia, Africa, Australia, New Zealand and the Americas (Peterson and Simon

1986; Vaughan and Geissler 2009). The gene centers for *D. carota* include Asia Minor, Transcaucasia, Iran, Turkmenistan, northwest India, Afghanistan, Tadjikistan, Uzbekistan and western Tian-shan mountain system of central Asia (Bradeen et al. 2002). Carrot domestication likely occurred in Central Asia (Iorizzo et al. 2013), but among Mediterranean regions, Tunisia is considered a center of biodiversity for *Daucus* and many other crops because of the diverse ecosystems and climatic conditions (Pottier Alapetite 1979; Le Floc'h et al. 2010), and is the center of diversity for important members of the *D. carota* clade, *D. carota* subsp. *capillifolius* and *D. sahariensis*. This clade is supported by a series of molecular studies (Spalik and Downie 2007; Arbizu et al. 2014a; Banasiak et al. 2016; Spooner et al. 2017), and is defined by shared chromosome numbers of $2n = 18$ and intercrosability of its component members (e.g., Krickl 1961; McCollum 1975, 1977; Vivek and Simon 1999; Hauser and Bjørn 2001), allowing for gene transfer of great potential economic use by carrot breeders.

The latest taxonomic monograph of *Daucus* by Sáenz Laín (1981) recognized 20 species, and Rubatzky et al. (1999) later estimated 25 species. In Tunisia, Pottier Alapetite (1979) recognized 11 *Daucus* species with several subspecies, to which we add *D. carota* subsp. *capillifolius*, described from western Libya, but found by us in adjacent eastern Tunisia. Many tools are now available for studying variability and the relationships among accessions including total seed protein, isozymes, molecular markers, and DNA sequence data (May 1992; Freville et al. 2001; Schlötterer 2004; Spooner et al. 2005). *Daucus* has been studied by a variety of DNA sequence techniques (Spalik and Downie 2007; Spooner et al. 2013, 2017; Arbizu et al. 2014a, 2016a, b; Banasiak et al. 2016). However, in addition to these powerful molecular methods to help define phylogenetic relationships, knowledge of the phenotype is critically important for practical taxonomic identifications and traits of use for plant breeders (Sudré et al. 2010). Although many floristic treatments have been published in the past few decades, many of *Daucus* species, and most critically for this study the subspecies variation in *D. carota*, are not fully understood. Identifications frequently use different characters and character states in their taxonomic keys and have incomplete synonymies which preclude comparison of their taxonomic

concepts and ambiguous identifications. Recently, attempts were made by Spooner et al. (2014) and Arbizu et al. (2014b) to quantify and describe the morphological variation in the *Daucus* collection conserved at the US genebank at the North Central Regional Plant Introduction Station (NCRPIS), for wild and cultivated carrots. We here examine the morphological diversity within the Tunisian National GeneBank collections of the *D. carota* clade at the inter- and intra-specific levels, using a larger morphological dataset than done previously with just fruit characters (Mezghani et al. 2014). We also expand the study of Mezghani et al. (2014) by assessing 32 morphological characters related to the full range of plant parts to include leaves, stems, and flowers, and in addition analyze the data from all individuals, not averaged data as in all prior studies. The result of this study adds to a growing body of morphological data in *Daucus* to aid in defining these taxa, will serve to produce an updated monograph of the genus, and will aid in the management, conservation and use of genetic resources.

Materials and methods

Plant material

We examined nine individuals each from 45 accessions (405 individuals in total) of the *Daucus carota* clade derived from a collection of 103 accessions conserved at the National Gene Bank of Tunisia (NGBT). The accessions were selected to maximize the diverse geographic and bioclimatic areas present in Tunisia, based on our previous study (Mezghani et al. 2014). We also added two newly collected accessions of *D. carota* subsp. *gummifer* from Galite Island, Tunisia. Geographic and bioclimatic data related to the accessions are shown in Table 1. Geographic Positioning System coordinates were imported into DIVA-GIS program version 7.5 (<http://www.diva-gis.org/>) and served as input data for mapping the accessions (Fig. 1).

Experimental field trial

The experiment was conducted in 2015 under field conditions at the High Institute of Agronomy of Chott Mariem (35.1182 N; 107297°E) using a randomized

block design, with 45 accessions, two replications and 5 plants per plot. Two 1-m row of each accession was directly and manually seeded in the field. Spacing was 3 m between blocks, 1.5 m between rows, and 0.2 m between plants. During culture, agronomic practices including irrigation, weeding, and fertilization were conducted uniformly as required in all plots.

Characters recorded

Daucus accessions were scored for 10 quantitative (continuous) and 22 qualitative (discontinuous) traits related to leaf, stem and flower (Table 2), using nine randomly selected plants per accession. The selection of characters was made following the International Board for Plant Genetic Resources guidelines for wild and cultivated carrot (*Daucus*) descriptors (IPGRI 1998), and adding additional characters used by Pottier Alapetite (1979) and Sáenz Laín (1981). Some of morphological characters were measured at the sampling location and others for microscopic observation in the laboratory. Quantitative traits were measured with a ruler or caliper while qualitative characters were based on scoring and coding according to the IPGRI (1998) descriptors. Images of different plant parts were made in the field and in the laboratory with a digital camera and images are available on the NGBT data base. Herbarium vouchers of the different species and subspecies are deposited at the NGBT herbarium.

Data analyses

Twenty-two of the 32 characters were scored and analyzed as discontinuous variables; the remaining ten were treated as continuous (Table 2). The operational taxonomic unit (OUT) was the individual, not the means of accessions as in our prior analyses (Mezghani et al. 2014; Spooner et al. 2014), precluding the need for assessing derivatives of averages or modes, and providing a larger and statistically more reliable data set.

For calculating the diversity parameters only, the overall entry mean value and the standard deviation were used to convert quantitative characters into qualitative ones (Jaradat et al. 2004) and frequencies were obtained from class intervals. The diversity was measured for each morphological character by using the standardized Shannon–Weaver (Shannon and Weaver

Table 1 Taxon name, origin, and bioclimatic zones of the *Daucus* populations studied here

Code	Original identification	Source (province, location)	Bioclimatic zone ^a
10449	<i>D. carota</i> subsp. <i>gummifer</i>	Bizerte, Galite Island	SH
10455	<i>D. carota</i> subsp. <i>gummifer</i>	Bizerte, Galite Island	SH
10820	<i>D. carota</i> subsp. <i>capillifolius</i>	Medenine, Djerba Island, Boumellel	ASW
10822	<i>D. carota</i> subsp. <i>capillifolius</i>	Medenine, Djerba Island, Houmet Souk	ASW
10823	<i>D. carota</i> subsp. <i>capillifolius</i>	Medenine, Djerba Island, between Ajim and Guellala	ASW
10833	<i>D. carota</i> subsp. <i>capillifolius</i>	Medenine, road between Zarzis and Djerba	ACW
10834	<i>D. carota</i> subsp. <i>capillifolius</i>	Medenine, road between Zarzis and Djerba	ACW
10835	<i>D. sahariensis</i>	Medenine, North of Bir Lahmar	ACW
10839	<i>D. carota</i> subsp. <i>sativus</i>	Tataouine, Farm of Ghrabi Family, north of Tataouine	S
10840	<i>D. carota</i> subsp. <i>sativus</i>	Tataouine, farm of Ghrabi Family, north of Tataouine	S
10842	<i>D. sahariensis</i>	Tataouine, road between Bir Lahmar and Tatounine	S
10843	<i>D. carota</i> subsp. <i>capillifolius</i>	Gabes, road from Medenine to Matmata	ASW
10845	<i>D. carota</i> subsp. <i>capillifolius</i>	Gabes, road between Matmata and Gabes	ASW
10850	<i>D. carota</i> subsp. <i>capillifolius</i>	Gafsa, road from El Guettar to Bou Omrane	ACW
10853	<i>D. carota</i> subsp. <i>capillifolius</i>	Gafsa, road from Gafsa to Bir El Hefay	ACW
10862	<i>D. carota</i> subsp. <i>capillifolius</i>	Sidi Bouzid, road from Sidi Bouzid to Regueb	ACW
10869	<i>D. carota</i> subsp. <i>capillifolius</i>	Sfax, Sakiat Eddayer	ASW
10871	<i>D. carota</i> subsp. <i>sativus</i>	Sfax, Ellouza Village	ASW
10872	<i>D. carota</i> subsp. <i>capillifolius</i>	Sfax, Ellouza Village	ASW
10873	<i>D. carota</i> subsp. <i>sativus</i>	Sfax, Ellouza Village	ASW
10876	<i>D. carota</i> subsp. <i>sativus</i>	Sfax, seed market in village of Hagez	ASW
10891	<i>D. carota</i> subsp. <i>capillifolius</i>	Sousse, road from Sousse to Tunis 27 km north of Sousse	SASW
10895	<i>D. carota</i> subsp. <i>capillifolius</i>	Sousse, road between Enfidha and Kairouan	SASW
10898	<i>D. carota</i> subsp. <i>capillifolius</i>	Sousse, road between Sousse and Hammamet	SASW
10905	<i>D. carota</i> subsp. <i>carota</i>	Zaghouan, road between Fahs and El Magren	SASW
10911	<i>D. carota</i> subsp. <i>carota</i>	Nabeul, road between Saliman and Sidi Aissa	SASW
10912	<i>D. carota</i> subsp. <i>carota</i>	Nabeul, road to Takelsa	SASW
10913	<i>D. carota</i> subsp. <i>carota</i>	Nabeul, road between Takelsa and El Haouaria	SASW
10918	<i>D. carota</i> subsp. <i>carota</i>	Ariana	SASW
10922	<i>D. carota</i> subsp. <i>carota</i>	Ariana, road to Bouhnach	SASW
10927	<i>D. carota</i> subsp. <i>carota</i>	Bizerte, south side of Ichkeul	SH
10928	<i>D. carota</i> subsp. <i>carota</i>	Bizerte, west of Ghezab	SH
10933	<i>D. carota</i> subsp. <i>carota</i>	Bizerte, Sajnane	SH
10935	<i>D. carota</i> subsp. <i>carota</i>	Beja, road from Sajnan to Tabarka	SH
10936	<i>D. carota</i> subsp. <i>carota</i>	Jendouba, Tabarka	H
10937	<i>D. carota</i> subsp. <i>carota</i>	Jendouba, Ain Draham	H
10943	<i>D. carota</i> subsp. <i>carota</i>	Jendouba, road from Aïn Draham to Beja	SH
10944	<i>D. carota</i> subsp. <i>maximus</i>	Jendouba, road from Ain Draham to Beja	SH
10947	<i>D. carota</i> subsp. <i>carota</i>	Beja, road between Beja and Tunis	SH
10951	<i>D. carota</i> subsp. <i>maximus</i>	Beja, road between Beja and Nefza	SH
10955	<i>D. carota</i> subsp. <i>carota</i>	Jendouba, Tabarka	SH
10956	<i>D. carota</i> subsp. <i>capillifolius</i>	Sousse, road between Sousse and Kairouan	SASW
10957	Putative hybrid of <i>D. carota</i> subsp. <i>capillifolius</i> × <i>D. carota</i>	Kairouan, road between Kairouan and Fahs	ACW

Table 1 continued

Code	Original identification	Source (province, location)	Bioclimatic zone ^a
10958	Putative hybrid of <i>D. carota</i> subsp. <i>capillifolius</i> × <i>D. carota</i>	Kairouan, road between Kairouan and Nassr Allah	ACW
10961	<i>D. carota</i> subsp. <i>capillifolius</i>	Kairouan, road between Bouhajla and Menzel Hached	ACW

^a Bioclimatic zones are defined according to Emberger's (1966) coefficient: ASW arid at soft winter, SAWS semi-arid at soft winter, ACW arid at cool winter, SACW semi-arid at cool winter, S Saharian

1949, as referred by Al Khanjari et al. 2008) diversity index, designed as H' ($H' = -\sum p_i (\log_2 p_i) / \log_2 n$, where p_i = frequency proportion of each descriptor state, n = number of states for each descriptor) was classified as high ($H' \geq 0.60$), intermediate ($0.40 \leq H' < 0.60$) or low ($0.10 \leq H' < 0.40$) as described by Eticha et al. (2005).

Multivariate analyses were conducted in JMP 10.0.0 software (SAS Institute 2012). We ran three types of analyses to explore the best ways to distinguish the accessions: (1) principal components analysis (PCA) using all of the data; (2) cluster analysis (average similarity, standardizing the data) using all of the data; (3) canonical discriminant analysis (CDA) and stepwise discriminant analysis (SDA) (linear, common covariance) using the ten continuous variables to obtain a model whose variables were significant in correctly identifying accession composition, with characters removed one at a time (if needed) until the model F-test p value was less or equal to 0.05. Scatter plots were then constructed to illustrate character-state distributions of the twelve best qualitative or quantitative characters distinguishing the taxa based on principal components and discriminate analyses.

Results

Diversity analysis

Large natural variation was found among accessions for the majority of traits (Table 2). Accessions were found homogenous only for four characters i.e., umbel type (UT), position of involucral bracts on primary umbel (PIB), anther color (AC) and symmetry of peripheral flowers (SPF). All accessions have compound umbels with deflexed involucral bracts, yellow anthers and symmetrical peripheral flowers. Estimated

diversity (H') for polymorphic traits ranged from 0.26 for foliage coverage (FoC) to 0.98 for flowering pattern within plants (FP) with overall means of 0.75, 0.77 and 0.76 for qualitative, quantitative and grand diversity mean, respectively. High phenotypic variability ($H' \geq 0.6$) was observed for all quantitative characters and 16 qualitative characters. Intermediate variation ($0.4 \leq H' < 0.6$) was observed for petiole shape in transverse section (PSTS) and stem ridging (SR). Low variation indicated the dominance of one character state over the others while high variation indicated equitable distribution of the different states as shown by frequency distribution.

Phenetic analyses

Both the PCA (Fig. 2), using all ten quantitative and 18 polymorphic qualitative characters, and CDA (Fig. 3), using the ten quantitative characters, defined four phenetic clusters: (1) *D. sahariensis*, (2) *D. carota* subsp. *capillifolius*, (3) *D. carota* subsp. *carota*, partially intergrading with *D. carota* subsp. *gummifer*, and (4) *D. carota* subsp. *sativus*, partially intergrading with putative hybrids between *D. carota* subsp. *capillifolius* and *D. carota* subsp. *carota*. While the CDA used only ten of the 28 total characters, the groups were still separated from each other. The 95% confidence ellipses about the means of the groups overlap only in cluster 4 containing putative hybrids between *D. carota* subsp. *capillifolius* and *D. carota* subsp. *carota* and *D. sativus*. PCA and CDA are both ordination techniques, but PCA makes no assumptions of group membership of OTUs. It attempts to portray multidimensional variation in the data set in the fewest possible dimensions, while maximizing the variation. CDA uses assigned groups to derive a linear combination of the variables (here morphological characters) that produces the greatest separation of the groups. As such, it tends to show slightly less overlap

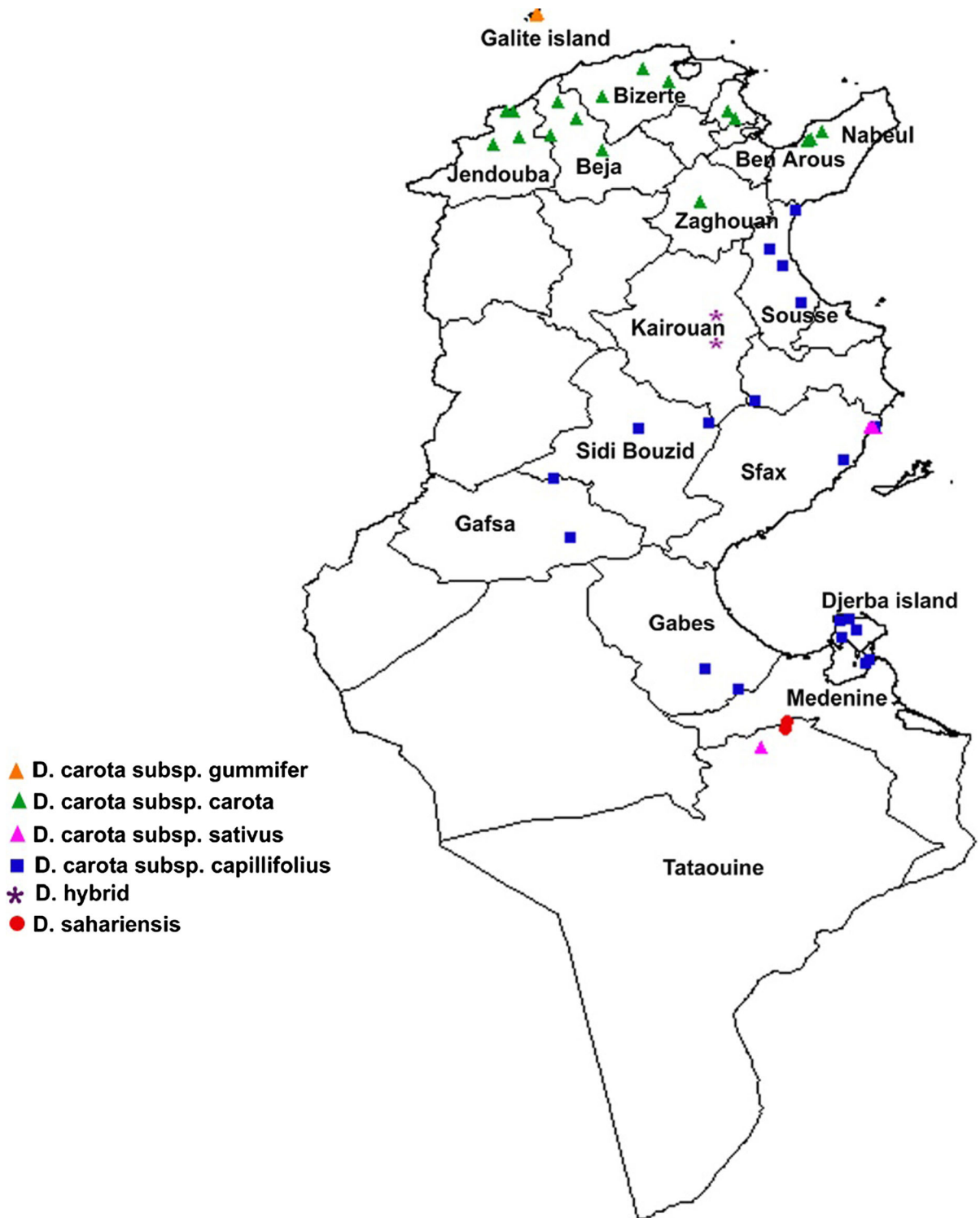


Fig. 1 Collection sites of the accessions examined in Tunisia

Table 2 Morphological descriptors, descriptor states, their codes for numerical analysis, frequency distribution, and diversity index of *Daucus* accessions listed in Table 1

Trait/descriptor	Descriptor acronym	Type	Descriptor state	Class	Frequency (%)	Diversity index (H')
<i>Leaf</i>						
Petiole shape in transverse section	PSTS	QL	Round	1	90.8	0.44
			Semi-round	2	9.2	
Petiole thickness (mm)	PT	QN	Narrow (≤ 2.7)	1	14.5	0.76
			Intermediate (2.8–4.3)	2	68.4	
			Wide (≥ 4.4)	3	17.1	
Anthocyanin coloration in petiole	ACP	QL	Uncolored	1	52.4	0.86
			Slightly colored	3	20.5	
			Intermediate	5	12.5	
			Strongly colored	7	14.6	
Petiole and leaf hairiness	PLH	QL	Sparse	1	45.7	0.80
			Intermediate	3	15.8	
			Strong	5	35.8	
			Very strong	7	2.7	
Length of primary basal leaflet (cm)	LPBL	QN	Short (≤ 4.8)	1	16.8	0.76
			Intermediate (4.9–9)	2	68.4	
			Elongated (≥ 9.1)	3	14.8	
Mature leaf length (cm)	MLL	QN	Short (≤ 16.5)	1	15.4	0.70
			Intermediate (16.6–26.5)	2	72.4	
			Elongated (≥ 26.6)	3	12.2	
Mature leaf width (cm)	MLW	QN	Narrow (≤ 8)	1	13.1	0.69
			Intermediate (8.1–14.7)	2	73.8	
			Wide (≥ 14.8)	3	13.1	
Leaf shape	LS	QL	Pinnatifid	2	4.4	0.74
			Bipinnate	4	37.8	
			Bi-tripinnate	5	57.8	
Leaf color	LC	QL	Yellow green	1	15.6	0.95
			Green	2	25.6	
			Grey-green	3	39.3	
			Purple green	4	19.5	
Leaf growth habit	LGH	QL	Prostrate	3	68.9	0.76
			Semi-erect	5	15.6	
			Erect	7	15.5	
Foliage coverage	FoC	QL	Sparse	3	4.4	0.26
			Strong	7	95.6	
Foliage width	FoW	QL	Narrow	3	12.1	0.73
			Intermediate	5	70.1	
			Wide	7	17.8	
<i>Stem</i>						
Stem color	SC	QL	Yellow green	1	15.6	0.89
			Green	2	32.8	
			Grey-green	3	42.2	

Table 2 continued

Trait/descriptor	Descriptor acronym	Type	Descriptor state	Class	Frequency (%)	Diversity index (H')
Stem ridging	SR	QL	Purple green	4	9.4	0.58
			Light	3	4.4	
			Intermediate	5	77.8	
Stem hairiness	SH	QL	Conspicuous	7	17.8	0.93
			Glabrous	1	42.2	
			Soft hairs	2	16.8	
Mean stem length (cm)	MSL	QN	Scabrous	3	41	0.79
			Short (≤ 59.4)	1	17.3	
			Intermediate (59.5–126.1)	2	66.9	
Mean stem diameter at the extremity (mm)	MSDE	QN	Elongated (≥ 126.2)	3	15.8	0.79
			Narrow (≤ 2.2)	1	16.5	
			Intermediate (2.2–4.1)	2	66	
Mean stem diameter at the base (mm)	MSDB	QN	Wide (≥ 4.2)	3	17.5	0.77
			Narrow (≤ 4.9)	1	16	
			Intermediate (4.9–11.2)	2	68.4	
<i>Flower</i>						
Flowering pattern within plants	FP	QL	Determinate	1	57.8	0.98
			Indeterminate	2	42.2	
Umbel type	UmT	QL	Simple	1	0	0
			Compound	2	100	
Primary umbel shape at flowering	PUSF	QL	Both	3	0	0.85
			Convex	1	16.2	
			Flat-topped	2	60.3	
Type of involucre bracts on primary umbel	TIB	QL	Concave	3	23.5	0.92
			Relatively branched bracts	1	15.8	
			Relatively unbranched bracts	2	46.4	
Position of involucre bracts on primary umbel	PIB	QL	Comparatively broad ultimate segments	3	37.8	0
			Deflexed	1	100	
			Not deflexed	2	0	
Corolla color	CC	QL	White	1	57.3	0.75
			Yellow	3	38.2	
			Pink	5	4.5	
Color of the central flower of the umbellet	CCFU	QL	White	1	14.9	0.83
			Yellow	3	38.2	
			Pink	5	4.5	
			Purple	6	42.4	
Symmetry of peripheral flower	SPF	QL	Symmetrical	1	100	0
			Asymmetrical	2		
Anther colour	AC	QL	Yellow	2	100	0
			Pink	4	0	

Table 2 continued

Trait/descriptor	Descriptor acronym	Type	Descriptor state	Class	Frequency (%)	Diversity index (H')
Primary umbel shape at fruit setting	PUSFS	QL	Not contracted	1	8.8	0.85
			Slightly contracted	2	14.4	
			Strongly contracted	3	76.8	
Density of flowers in umbels	DFU	QL	Low	3	42.2	0.97
			Intermediate	5	35.5	
			High	7	22.3	
Total number of umbels per plant	TNUP	QN	Low (≤ 7.3)	1	18	0.82
			Intermediate (7.4–22.5)	2	63.7	
			High (≥ 22.6)	3	18.3	
Width of primary open umbel (cm)	WPOU	QN	Narrow (≤ 7.7)	1	15.8	0.78
			Intermediate (7.8–12.7)	2	66.7	
			Wide (≥ 12.8)	3	17.5	
Average number of umbellets per umbel	ANUU	QN	Low (≤ 29.3)	1	16.5	0.81
			Intermediate (29.4–89)	2	64.7	
			High (≥ 89.1)	3	18.8	

^a Quantitative characters were converted to phenotypic classes with the class boundaries as described by Jaradat et al. (2004)

of the two taxa in group 3, *D. carota* subsp. *carota* and *D. carota* subsp. *gummifer*.

The cluster analysis (Fig. 4), however, defined an additional two groups (now six in total), splitting

group 3 above into two, *D. carota* subsp. *carota* and *D. carota* subsp. *gummifer*, and group 4 into two, *D. carota* subsp. *sativus*, and the putative hybrids between *D. carota* subsp. *capillifolius* and *D. carota* subsp. *carota*. While the PCA grouped the taxa separately, it did not group the nine individuals per accession together, that is, the individuals were largely intermixed among the different accessions among the larger taxon cluster. Like the PCA, the cluster analysis makes no assumptions of group membership of OTUs. These cluster analyses and CDA results were not significantly different, however, because in both cases where the extra groups are split, they are barely distinguished on the cluster analysis (Fig. 4).

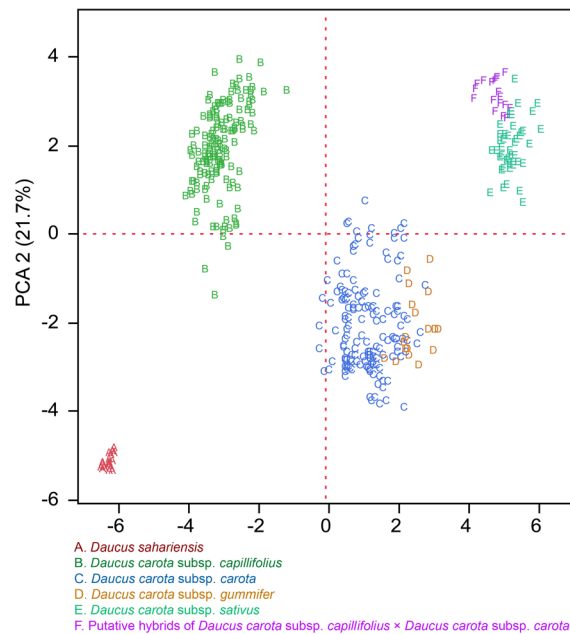


Fig. 2 Principal components analysis of the accessions examined in Tunisia

Character state distributions

All ten of the continuously variable characters were significant ($p < 0.001$) to discriminate the species based on the discriminate analysis, listed from their most to the least as determined by the F ratio statistic: mean stem length (MSL), average number of umbellets per umbel (ANUU), mean stem diameter at the base (MSDB), mean stem diameter at the extremity (MSDE), total number of umbels per plant (TNUP), mature leaf length (MLL), width of primary open umbel (WPOU), petiole

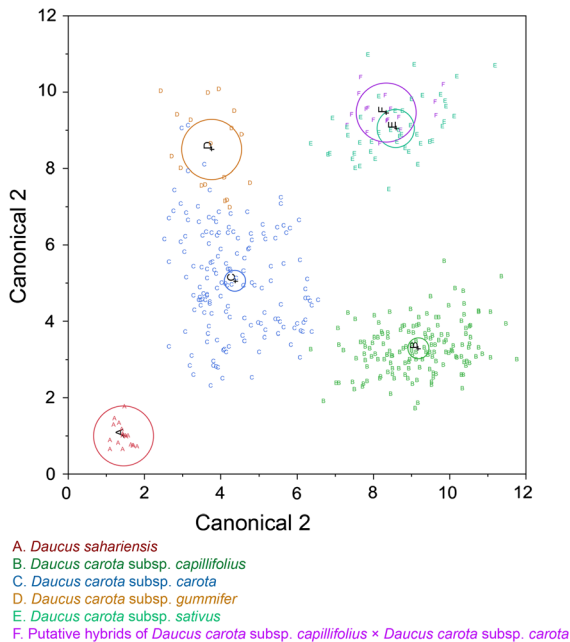


Fig. 3 Canonical discriminant analysis of the accessions examined here. *Ellipses* show the 95% confidence ellipse of each mean

thickness (PT), length of primary basal leaflet (LPBL), and mature leaf width (MLW). Based on high eigen-vectors (positive or negative) the following five characters were most important to distinguish the species in principle component 1, listed from their highest to lowest: ANUU (above), density of flowers in umbels (DFU), corolla color (CC), type of involucre bracts on primary umbel (TIB), and leaf shape (LS); and in principle component 2: MSL (above), TNUP (above), anthocyanin coloration in petiole (ACP), petiole and leaf hairiness (PLH), and MLL (above). Four of these characters are repeated (ANUU, MLL, MSL, TNUP), making for 16 quantitative or qualitative characters best distinguishing the taxa analyzed here. Based on ranking the F statistic of the 10 characters from the discriminant analysis, and inspection of character state distributions from the 10 characters as determined by the factor loadings PCA analysis, the best 12 of these 16 characters (eight quantitative and four qualitative) are presented in Fig. 5. These distributions reflect the phenetic results of the PCA and CDA. That is, *D. sahariensis* is most clearly defined, as is reflected in six of the 12 characters shown in Fig. 5 that have no or very little overlap of characters with any of the other taxa: mean stem length, mean stem diameter at base, average number of

umbellets per umbel, mature leaf length, mature leaf width, and leaf shape; easily distinguishing this species as a relatively low-growing plant with small stems, few umbellets/umbel, and small pinnatifid leaves.

Daucus carota subsp. *capillifolius* (ignoring the putative hybrids) is also well distinguished by its corolla color, leaf shape, and type of involucre bract on primary umbel; essentially easily defining this subspecies as having yellow corollas, bipinnate leaves, and comparatively broad ultimate segments of the involucre bracts on the primary umbel.

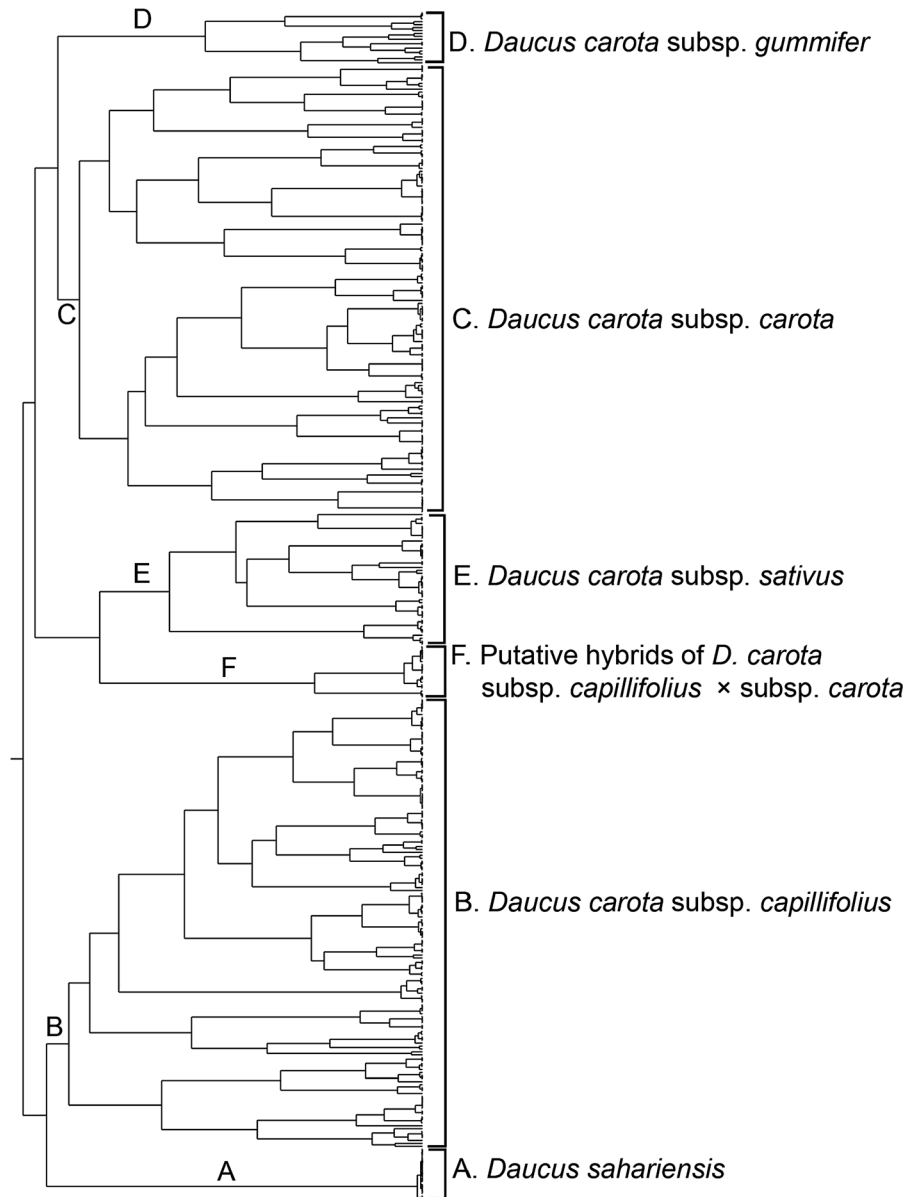
Distinguishing *D. carota* subsp. *carota* and *D. carota* subsp. *gummifer* (as a group) from the other taxa, however, is more difficult and relies on a series of partially overlapping character states (polythetic support). These include mean stem length, corolla color, type of involucre bract on primary umbel, and leaf shape. Distinguishing *D. carota* subsp. *carota* and *D. carota* subsp. *gummifer* from each other is more difficult. Again, overlapping characters are used, to include shorter stems and more umbellets/umbel. *Daucus carota* subsp. *gummifer* (sensu lato) is a highly polymorphic taxon distributed very near the western Mediterranean coasts and the Atlantic coasts of UK, France, Spain and Portugal, and also is distinguished by characters difficult to measure here, such as stiff parts in the peduncles and inflorescences, and often thickened and shiny (varnished) leaves.

Finally, distinguishing *D. carota* subsp. *sativus* (cultivated carrot) and the putative hybrids between *D. carota* subsp. *capillifolius* and *D. carota* subsp. *carota* (as a group) from the other taxa also is difficult and relies on a series of partially overlapping character states. These include mean stem diameter at the base and relatively branched bracts on the primary umbel.

Discussion and conclusions

Information on diversity and relationships within and among crop species and their wild relatives is essential for the efficient utilization of plant genetic resource collections and presents an efficient proxy of taxonomic relationships (Sun and Wong 2001; Drzewiecki et al. 2003). In breeding programs, characterization of accessions based on multiple traits can be used as a management tool in regenerations to allow validating the identity of an accession. Evaluation data are used when searching the gene bank for useful germplasm

Fig. 4 Phenogram of the accessions examined in Tunisia



(DeLacy et al. 2000). International surveys clearly show the need to save and manage the local germplasm of each country, since accessions may contain valuable genes of biotic and abiotic stress for crop breeding (Mengistu et al. 2015).

For *Daucus*, the taxonomy of the members of the *Daucus carota* clade remains unresolved (Arbizu et al. 2014b, 2016b). The morphological characters used in our study revealed considerable diversity within a local (Tunisian) *Daucus* collection maintained at the National Gene Bank of Tunisia as shown by mean

diversity indexes of 0.75, 0.77 and 0.76 for qualitative, quantitative and total diversity respectively, showing Tunisia to contain significant diversity for *Daucus* in the southern Mediterranean region. Except for *D. sahariensis* and *D. carota* subsp. *capillifolius*, there were no taxon-specific quantitative or qualitative characters, but rather taxonomic differentiation relies on a series of partially overlapping character states (polythetic support). This is reflective of their very designation as subspecies rather than species, by the continuing disagreement of taxonomic boundaries,

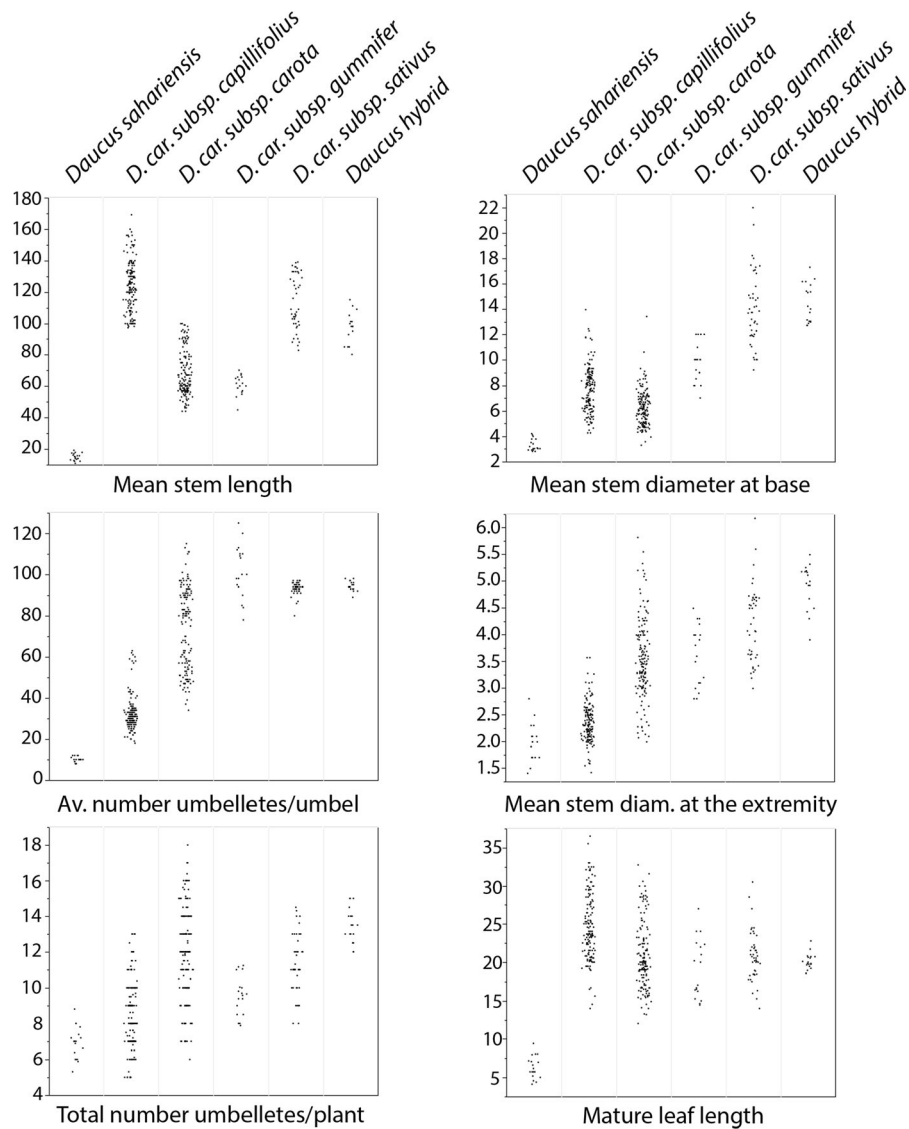


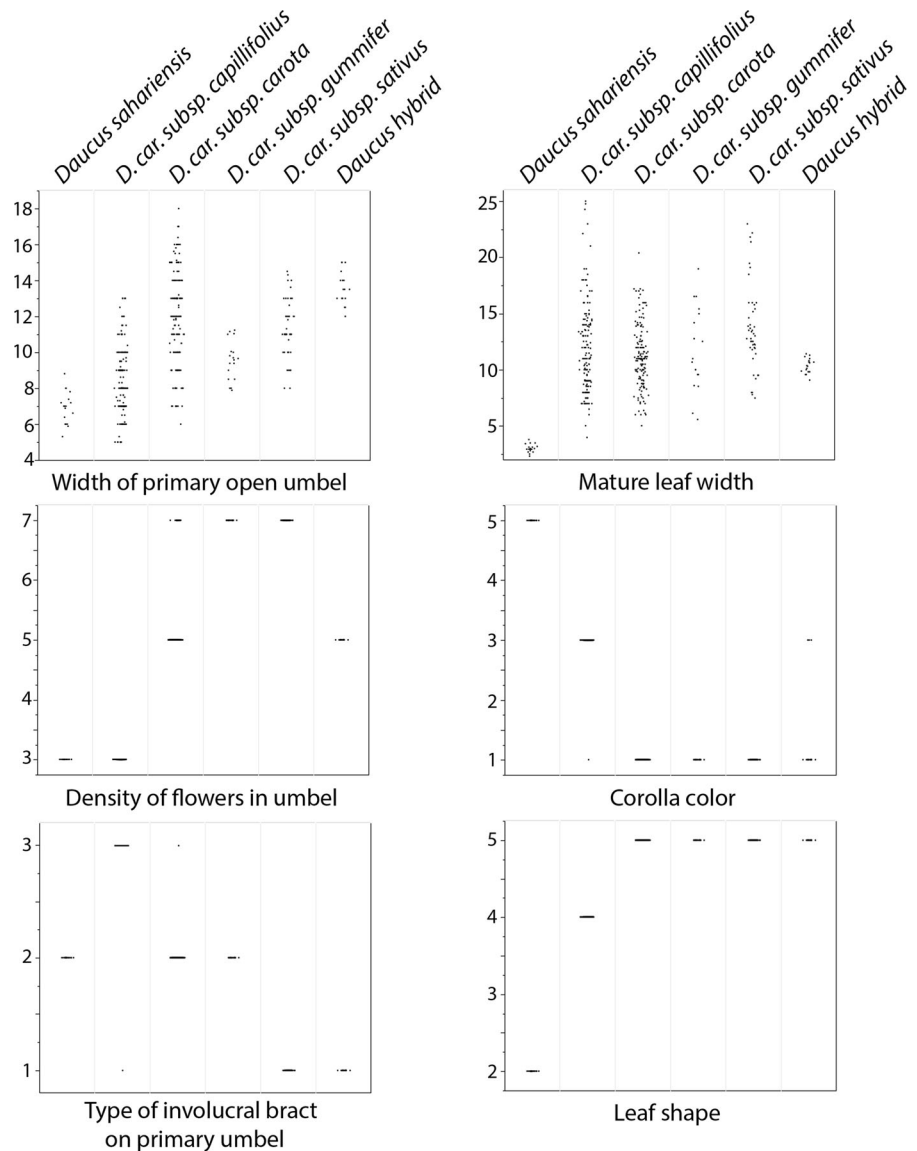
Fig. 5 Character state distributions of the 12 best quantitative and qualitative characters distinguishing the taxa examined in Tunisia

and by different use of characters by taxonomists to differentiate them. These taxonomic problems are possibly the result of their ability to freely exchange genes in nature and experimentally (above). Our study, using a well-dispersed and comprehensive germplasm collecting in a local area containing much of this diversity (Tunisia), adds to a growing body of evidence documenting the poor differentiation of these taxa.

Two Tunisian accessions we examined (10944 and 10951) possess characters of yet another possible

subspecies, *D. carota* subsp. *maximus* (Desf.) Ball: inflated, leathery stipule bases, relatively large umbels, and fruits with spines shorter than the diameter of the fruit and stellate at the top (Mezghani et al. 2014). Using morphological data, Spooner et al. (2014) suggested that only two subspecies of *D. carota* can be distinguished: subsp. *carota* and subsp. *gummifer* corresponding to the two species (*D. carota* and *D. gingidium*) recognized by Onno (1937) and Pottier Alapetite (1979) or to the two “species aggregates” or “subgroups”, recognized by Small (1978) and

Fig. 5 continued



Ruduron (2007). On the basis of molecular and morphological data Arbizu et al. (2014a, b) later lowered *D. capillifolius* to subspecies rank of *D. carota*. Tavares et al. (2014) attempted to distinguish the subspecies of *D. carota* native to Portugal and supported subsp. *maximus* as a distinct taxon from other taxa by morphometric analysis of the fruits and chemical characterization of essential oils, concluding as Saenz de Rivas and Heywood (1974) that subsp. *maximus* should be considered a species (i.e., *D. maximus*) rather than a subspecies of *D. carota*. A broader molecular (genotyping by sequencing) study by Arbizu et al. (2016b) did not support the subspecies

of *D. carota* to form monophyletic lineages, except subsp. *capillifolius* which is an apespecies occurring within the clade containing subsp. *carota*; that is, subsp. *maximus* and subsp. *gummifer* occurring in different clades containing geographically isolated populations of subsp. *carota* in Europe (France, Portugal and Italy) and subsp. *maximus* in Northern Africa (Morocco).

PCA and CDA results unexpectedly placed the putative hybrids between *D. carota* subsp. *capillifolius* and *D. carota* subsp. *carota* with *D. carota* subsp. *sativus*. We have no explanation for this other than knowledge that hybrids often exhibit a mosaic of

intermediate, parental, transgressive, and novel characters (Rieseberg and Ellstrand 1993), and the morphological results shown here do not reflect true hybrid origins. Alternatively, our hypothesis of hybridity may be incorrect and perhaps these are true hybrids with subsp. *sativus*.

The ultimate resolution of the taxonomic boundaries of members of the *D. carota* clade will require analysis of a broader range of accessions with additional molecular data, as we are pursuing collaboratively with a broader research group.

Compliance with ethical standards

Conflict of interest None of the authors have any financial or personal relationship that could inappropriately influence the work submitted for publication.

References

- Agrawal RC, Behera D, Saxena S (2007) Genebank information management system (GBIMS). *Comput Electron Agr* 59:90–96
- Al Khanjari S, Filatenko AA, Hammer K, Buerkert A (2008) Morphological spike diversity of Omani wheat. *Genet Resour Crop Evol* 55:1185–1195
- Arbizu C, Ruess H, Senalik D, Simon P, Spooner DM (2014a) Phylogenomics of the carrot genus (*Daucus*, Apiaceae). *Am J Bot* 101:1666–1685
- Arbizu C, Reitsma KR, Simon PW, Spooner DM et al (2014b) Morphometrics of *Daucus* (Apiaceae): a counterpart to a phylogenomic study. *Am J Bot* 101:2005–2016
- Arbizu C, Simon PW, Martínez-Flores F, Ruess H, Crespo MB, Spooner DM (2016a) Integrated molecular and morphological studies of the *Daucus guttatus* complex (Apiaceae). *Syst Bot* 41:479–492
- Arbizu CI, Ellison SL, Senalik D, Simon PW, Spooner DM (2016b) Genotyping-by-sequencing provides the discriminating power to investigate the subspecies of *Daucus carota* (Apiaceae). *BMC Evol Biol* 16:234
- Banasiak Ł, Wojewódzka A, Baczyński J, Reduron J-P, Piwczyński M, Kurzyńska-Młynik R, Gutaker R, Czarnocka-Cieciura A, Kosmala-Grzechnik S, Spalik K (2016) Phylogeny of Apiaceae subtribe Daucinae and the taxonomic delineation of its genera. *Taxon* 65:563–585
- Bradeen JM, Bach IC, Briard M, le Clerc V, Grzebelus D, Senalik DA, Simon PW (2002) Molecular diversity analysis of cultivated carrot (*Daucus carota* L.) and wild *Daucus* populations reveals a genetically nonstructured composition. *J Am Soc Hortic Sci* 127:383–391
- DeLacy IH, Skovmand B, Huerta J (2000) Characterization of Mexican wheat landraces using agronomically useful attributes. *Genet Resour Crop Evol* 47:591–602
- Drzewiecki J, Delgado-Licon E, Haruenkit R, Pawelzik E, Martin-Belloso O, Park YS, Jung ST, Trakhtenberg S, Gorinstein S (2003) Identification and differences of total proteins and their soluble fractions in some pseudocereals based on electrophoretic patterns. *J Agr Food Chem* 51:7798–7804
- Eticha F, Bekele E, Belay G, Börner A (2005) Phenotypic diversity in durum wheat collected from Bale and Wello regions of Ethiopia. *Plant Genet Resour* 3(1):35–43
- FAO (1997) The state of the world's plant genetic resources for food and agriculture. Food and Agriculture Organization of the United Nations, Rome
- Freville H, Justy F, Olivieri I (2001) Comparative allozyme and microsatellite population structure in a narrow endemic plant species, *Centaurea corymbosa* Pourret (Asteraceae). *Mol Ecol* 10:879–889
- Hauser TP, Bjørn GK (2001) Hybrids between wild and cultivated carrots in Danish carrot fields. *Genet Resour Crop Evol* 48:499–506
- Iorizzo M, Senalik DA, Ellison SL, Grzebelus D, Cavagnaro PF, Allender C, Brunet J, Spooner DM, Van Deynze A, Simon PW (2013) Genetic structure and domestication of carrot (*Daucus carota* subsp. *sativus* L.) (Apiaceae). *Am J Bot* 100:930–938
- IPGRI (1998) Descriptors for wild and cultivated carrot (*Daucus carota* L.). International Plant Genetic Resources Institute, Rome
- Jaradat AA, Shahid M, Al Maskri AY (2004) Genetic diversity in the Batini barley landrace from Oman: I. Spike and seed quantitative and qualitative traits. *Crop Sci* 44:304–315
- Krickl M (1961) Karotten: zur Frage der Verkreuzung mit der wilden Karotte. *Saatgut-Wirtschaft* 13:135–136
- Le Floch E, Boulos L, Vela E (2010) Catalogue synonymique commenté de la flore de Tunisie. Banque Nationale des Gènes de la Tunisie, Tunis
- May B (1992) Starch gel electrophoresis of allozymes. In: Hoelzel AR (ed) *Molecular genetic analysis of populations: a practical approach*. Oxford University Press, Oxford, pp 1–27
- McCollum GD (1975) Interspecific hybrid *Daucus carota* × *D. capillifolius*. *Bot Gaz* (Chicago, Ill) 136:201–206
- McCollum GD (1977) Hybrids of *Daucus gingidium* with cultivated carrots (*D. carota* subsp. *sativus*) and *D. capillifolius*. *Bot Gaz* (Chicago, Ill) 138:56–63
- Mengistu DK, Kiros AY, Pè ME (2015) Phenotypic diversity in Ethiopian durum wheat (*Triticum turgidum* var. *durum*) landraces. *Crop J* 3:190–199
- Mezghani N, Zaouali I, Bel Amri W, Rouz S, Simon PW, Hannachi C, Ghrabi Z, Neffati M, Bouzbida B, Spooner DM (2014) Fruit morphological descriptors as a tool for discrimination of *Daucus* L. germplasm. *Genet Resour Crop Evol* 61:499–510
- Onno M (1937) Die Wildformen von *Daucus* sect. *carota*. *Beih Bot Zentralbl* 56(B):83–136
- Peterson CE, Simon PW (1986) Carrot Breeding. In: Basset MJ (ed) *Breeding vegetable crops*. AVI Publishing Company, Westport, pp 321–356
- Pottier Alapetite G (1979) *Daucus*. Flore de la Tunisie, Angiospermes-Dicotyledones, Apetales, Dialypetales. Imprimerie Officielle de la République Tunisienne, Tunis, pp 615–621
- Rao NK, Hanson J, Dulloo ME, Ghosh K, Nowell D, Larinde M (2006) Manuel de manipulation des semences dans les

- banques de gènes. Manuels pour les banques de gènes No. 8. Bioversity International, Rome
- Rieseberg LH, Ellstrand NC (1993) What can molecular and morphological markers tell us about plant hybridization? *Crit Rev Pl Sci* 12:213–241
- Rubatzky VE, Quiros CF, Simon PW (1999) Carrots and related vegetable Umbelliferae. CABI Publishing, New York
- Ruduron J-P (2007) Ombellifères de France, vol 2. Société botanique du centre Ouest, Nercillac, p 564
- Saenz de Rivas C, Heywood VH (1974) Estudio preliminar sobre los *Daucus* en España peninsular. *An Inst Bot Cavanilles* 31:97–118
- Sáenz Laín C (1981) Research on *Daucus* L. (*Umbelliferae*). Actas III Congr ÓPTIMA. *An Jard Bot Madr* 37:481–534
- SAS Institute (2012) JMP, version 10.0.0. SAS Inst., Cary, NC
- Schlötterer C (2004) The evolution of molecular markers—just a matter of fashion? *Nat Rev Genet* 5:63–69
- Shannon CE, Weaver W (1949) The mathematical theory of communication. University of Illinois Press, Urbana
- Small E (1978) A numerical taxonomic analysis of the *Daucus carota* complex. *Canad J Bot* 56:248–276
- Spalik K, Downie SR (2007) Intercontinental disjunctions in *Cryptotaenia* (Apiaceae, Oenantheae): an appraisal using molecular data. *J Biogeogr* 34:2039–2054
- Spooner DM, van Treuren R, de Vicente MC (2005) Molecular markers for genebank management. IPGRI technical bulletin no. 10. International Plant Genetic Resources Institute, Rome
- Spooner DM, Rojas P, Bonierbale M, Mueller LA, Srivastav M, Senalik D, Simon PW (2013) Molecular phylogeny of *Daucus* (Apiaceae). *Syst Bot* 38:850–857
- Spooner DM, Widrechner MP, Reitsma KR, Palmquist DE, Grabi-Gammar Z, Neffati M, Bouzbida B, Ouabbou H, El Koudrim M, Simon PW (2014) Reassessment of practical species identifications of the USDA *Daucus carota* germplasm collection: morphological data. *Crop Sci* 54:1–13
- Spooner DM, Ruess H, Iorizzo M, Senalik D, Simon PW (2017) Entire plastid phylogeny of the carrot genus (*Daucus*, Apiaceae); concordance to nuclear data and mitochondrial and nuclear DNA insertions to the plastid. *Am J Bot* 104(2):296–312
- Sudré CP, Gonçalves LSA, Rodrigues R, do Amaral Júnior AT, Riva-Souza EM, Bento Cdos S (2010) Genetic variability in domesticated *Capsicum* ssp. as assessed by morphological and agronomic data in mixed statistical analysis. *Genet Mol Res* 9:283–294
- Sun M, Wong KC (2001) Genetic structure of three orchid species with contrasting breeding systems using RAPD and allozyme markers. *Am J Bot* 88:2180–2188
- Tavares AC, Loureiro J, Castro S, Coutinho AP, Paiva J, Cavaleiro C, Salgueiro L, Canhoto JM (2014) Assessment of *Daucus carota* L. (Apiaceae) subspecies by chemotaxonomic and DNA content analyses. *Biochem Syst Ecol* 55:222–230
- Teklu Y, Hammer K, Huang X, Röder M (2006) Analysis of microsatellite diversity in Ethiopian durum wheat landraces. *Genet Resour Crop Evol* 53:1115–1126
- Vaughan JG, Geissler CA (2009) The new Oxford book of food plants. Oxford University Press, Oxford
- Vivek BS, Simon P (1999) Phylogeny and relationships in *Daucus* based on restriction fragment length polymorphisms (RFLPs) of the chloroplast and mitochondrial genomes. *Euphytica* 105:183–189